



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

**Fatores associados ao desempenho escolar em
Matemática do SAEB 2013 para alunos do 3º ano do
Ensino Médio no Distrito Federal:**

Uma Aplicação de Regressão Multinível

Andressa de Lima Teixeira

Brasília

2015

Andressa de Lima Teixeira

**Fatores associados ao desempenho escolar em
Matemática do SAEB 2013 para alunos do 3º ano do
Ensino Médio no Distrito Federal:**

Uma Aplicação de Regressão Multinível

Trabalho de Conclusão de Curso apresentado ao
Departamento de Estatística da Universidade de
Brasília, como parte dos requisitos para a obtenção
do título de Bacharel em Estatística.

Orientador: Prof^a. **Maria Teresa Leão Costa**

Brasília

2015

Sumário

Introdução	1
Objetivos	3
1 Regressão Multinível	5
1.1 Introdução	5
1.2 Formulação do Modelo e Interpretação dos Parâmetros	6
1.3 Estimação dos Parâmetros	9
1.4 Intervalos de Confiança	10
1.5 Comparação de Modelos	11
1.6 Seleção de Variáveis	12
1.7 Coeficiente de Correlação Intraclasses	14
1.8 Coeficiente de Determinação R^2	15
1.9 Análise do Resíduo	16
2 Material e Método	19
2.1 Introdução	19
2.2 Material	20
2.3 Método	21
3 Resultados e Discussões	23
3.1 Perfil dos Alunos	23
3.2 Perfil das Escolas	27
3.3 Identificação de Possíveis Variáveis Explicativas	29
3.4 Modelagem	33

4 Conclusão	43
Referências Bibliográficas	48

Lista de Figuras

1.1	Estrutura Hierárquica	5
1.2	Estrutura dos dados para um modelo multinível com dois níveis	6
1.3	Gráfico dos Resíduos Padronizados versus escores normais	17
1.4	Gráfico dos resíduos versus a Variável Preditora	17
3.1	Distribuição dos alunos por dependência administrativa das escolas	28
3.2	Distribuição das notas em Matemática por sexo, raça, nível socioeconômico e se o aluno trabalha	31
3.3	Distribuição das notas em Matemática considerando se o aluno já reprovou, abandonou os estudos, concluiu o ensino fundamental no EJA e se faz o dever de casa de matemática	32
3.4	Distribuição das notas em Matemática pelas características da escola	33
3.5	Gráfico resíduos padronizados vs. valores preditos para o modelo final	39
3.6	Gráfico Quantil-Quantil para o modelo final	40

Lista de Tabelas

2.1	População de referência no Brasil do SAEB 2013	20
2.2	Renda média bruta familiar no mês em Reais por classe	21
3.1	Distribuição de frequência dos alunos do 3º ano do ensino médio do Distrito Federal por idade	24
3.2	Distribuição de frequência do sexo e raça dos alunos do 3º ano do ensino médio do Distrito Federal por dependência administrativa da escola	24
3.3	Distribuição de frequência da escolaridade das mães dos alunos do 3º ano do ensino médio do Distrito Federal por dependência administrativa da escola	25
3.4	Distribuição de frequência da escolaridade dos pais dos alunos do 3º ano do ensino médio do Distrito Federal por dependência administrativa da escola	25
3.5	Distribuição de frequência do nível socioeconômico dos alunos do 3º ano do ensino médio do Distrito Federal por dependência administrativa da escola	26
3.6	Distribuição de frequências da participação dos pais na vida escolar dos alunos do 3º ano do ensino médio do Distrito Federal por dependência administrativa da escola	26
3.7	Distribuição de frequências do histórico escolar dos alunos do 3º ano do ensino médio do Distrito Federal por dependência administrativa da escola	27
3.8	Distribuição de frequência para os hábitos de estudo dos alunos do 3º ano do ensino médio do Distrito Federal e dos professores por dependência administrativa da escola	27
3.9	Distribuição de frequências do professor por dependência administrativa da escola	29

3.10 Média do rendimento na prova de matemática do SAEB 2013 para o 3º ano do Ensino Médio no Distrito Federal segundo características dos alunos e Estatísticas dos teste de comparação	30
3.11 Características das escolas e Estatísticas dos teste do SAEB 2013 para o 3º ano do Ensino Médio	31
3.12 Modelo Nulo	34
3.13 Modelo com todas as variáveis explicativas do primeiro nível	35
3.14 Modelo com acréscimo das variáveis do segundo nível	36
3.15 Modelo com as variáveis significativas nos dois níveis	37
3.16 Modelo com acréscimo dos efeitos aleatórios	38
3.17 Teste χ^2 dos modelos	38

Resumo

Em estudos educacionais é importante considerar a estrutura hierárquica da população, visto que os alunos estão inseridos em escolas que podem desempenhar um papel relevante no rendimento escolar dos estudantes. Nesses casos uma técnica adequada é a regressão multinível, pois leva em consideração a dependência dos dados e é possível analisar todos os níveis da hierarquia.

Nesse trabalho foi construído um modelo multinível para o desempenho dos alunos do 3º ano do Ensino Médio no Distrito Federal na prova de matemática do SAEB no ano 2013.

Palavras-chave: Regressão Multinível, Modelos Hierárquicos, Avaliação Educacional, SAEB, Ensino Médio.

Introdução e Justificativa

No estudo de fenômenos em diferentes áreas deseja-se, muitas vezes, explicar o comportamento de uma variável (resposta) em função de um conjunto de variáveis explicativas. Com o intuito de representar essa relação é importante estabelecer um modelo estatístico. Essa técnica é chamada de regressão.

Em particular, no campo educacional é de interesse, por exemplo, analisar o desempenho dos estudantes em determinadas disciplinas e verificar quais fatores estão associados a esse desempenho. Sendo assim, seleciona-se inicialmente uma amostra aleatória de observações na escola (nível superior) e a partir dessas amostras seleciona-se também de forma aleatória amostras dos alunos (nível inferior). Dessa maneira, os alunos estão agrupados em escolas.

Nesse caso não podemos supor independência entre as observações visto que a convivência dos alunos em suas respectivas escolas faz com que o grupo de alunos se assemelhe cada vez mais e aumente a diferença em relação a outras escolas.

A correlação entre os alunos da mesma escola tende ser mais alta do que alunos em escolas distintas, existindo então uma correlação intraclasse, o que torna o modelo de regressão tradicional inadequado pois esse não leva em consideração a dependência entre os indivíduos do mesmo nível. Portanto o método tradicional de regressão pode levar à altas estimativas dos erros padrão, o que leva a "resultados significativos" nos testes de hipóteses na maioria dos casos.

Em situações como essa torna-se mais interessante utilizar a Regressão Multi-nível, uma vez que essa é utilizada quando a estrutura dos dados é hierárquica, ou seja, dados agrupados em diferentes níveis. Uma das vantagens de se utilizar a Regressão Multi-nível quando se tem dados hierárquicos é corrigir a superestimação dos erros padrão,

além de se tornar mais precisos os intervalos de confiança e testes de hipóteses.

Na Regressão Multinível as variáveis podem ser definidas em cada nível. As relacionadas ao aluno encontram-se, por exemplo, no nível 1 como sexo e local de residência, e as relacionadas a escola no nível 2, como infraestrutura.

Objetivos

Objetivo Geral

Este projeto tem como objetivo estudar regressão multinível bem como aplicar os conhecimentos adquiridos na construção de um modelo que identifique os fatores relacionados ao desempenho dos estudantes do 3º ano do Ensino Médio do Distrito Federal na prova de Matemática do SAEB 2013.

Objetivos Específicos

- Estudar Modelos Multiníveis;
- Estudar a linguagem do *software* estatístico R para aplicar a metodologia estudada;
- Organizar o Banco de dados da aplicação;
- Aplicar a modelagem multinível no estudo dos fatores que influenciam o desempenho dos alunos do 3º ano do Ensino Médio na prova do SAEB 2013.

Capítulo 1

Regressão Multinível

1.1 Introdução

Estudos em diversas áreas tais como ciências humanas e sociais se baseiam em dados que estão estruturados de forma hierárquica, como por exemplo, pacientes dentro de hospitais ou alunos inseridos em turmas que por sua vez estão agrupados em escolas.

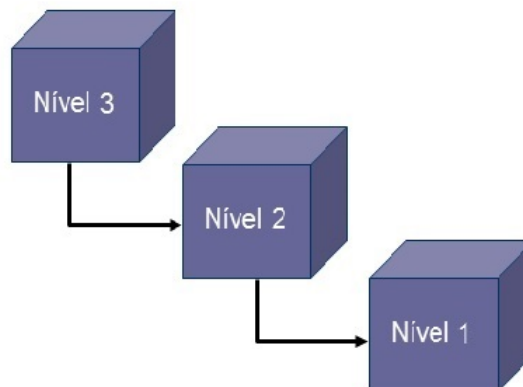


Figura 1.1: Estrutura Hierárquica

Em casos assim o modelo de regressão tradicional torna-se inadequado, pois a suposição de independência é violada visto que existe uma correlação entre indivíduos pertencentes a um mesmo nível. Consequentemente, o modelo de regressão tradicional pode levar a altas estimativas do erro padrão, o que influenciará na parte inferencial do modelo, tais como super estimativas do Intervalo de Confiança e Testes de Hipóteses.

1.2 Formulação do Modelo e Interpretação dos Parâmetros

Considere, por exemplo, um estudo que deseja identificar fatores associados ao desempenho dos alunos em uma Prova de Língua Portuguesa. A amostra utilizada é composta por K turmas com n_k alunos cada uma. A variável resposta Y "Desempenho do aluno na prova de português" está associada ao nível inferior (alunos). Nesse caso as variáveis explicativas podem ser definidas em cada nível hierárquico. No nível alunos (inferior) existem duas variáveis explicativas, sendo a primeira sexo (X_1 : 0 quando feminino e 1 quando masculino) e Leitura (X_2 : Quantidade de livros lidos nos últimos três meses). No nível da turma (superior) a variável explicativa é dada pela experiência do professor de português (Z : Anos de experiência do professor).

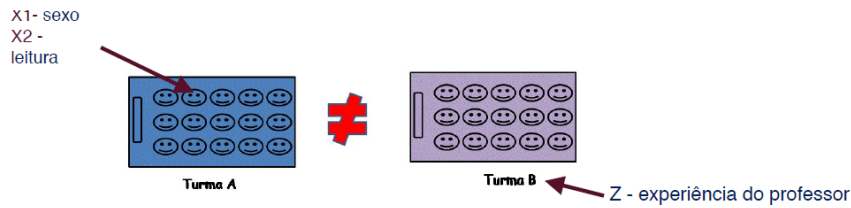


Figura 1.2: Estrutura dos dados para um modelo multinível com dois níveis

As variáveis do exemplo são mensuradas de forma direta em cada nível. Em outros casos se pode mover as variáveis de um nível para outro por agregação, ou seja, quando a variável do nível inferior é realocada para o nível superior. No exemplo poderia ser atribuída à escola a média da leitura de seus respectivos alunos. Ou por desagregação, quando a variável do nível superior é realocada para o nível inferior, ou seja, poderia ser atribuído a cada aluno o respectivo tempo de experiência do professor.

Para análise dos dados pode-se construir uma equação de regressão para cada uma das K turmas afim de prever a variável resposta Y .

$$Y_{ik} = \beta_{0k} + \beta_{1k}X_{1ik} + \beta_{2k}X_{2ik} + e_{ik}. \quad (1.1)$$

em que $k = 1, \dots, K$ é o índice associado a turma e $i = 1, \dots, n_k$ é o índice associado ao nível inferior e $e_{ik} \sim N(0, \sigma_e^2)$.

Como pode ser observado na equação 1.1, em regressão multinível os coefici-

entes β_{0k} , β_{1k} e β_{2k} são variáveis aleatórias que variam de acordo com a turma. Nesse caso assume-se que os coeficientes β_{ik} tem distribuição normal multivariada.

A interpretação dos parâmetros segue de forma análoga ao modelo de regressão tradicional. Em geral, quanto maior o valor do intercepto maior será, em média, o desempenho dos alunos em português. Turmas com valores elevados do coeficiente β_{1k} indicam que a diferença entre meninos e meninas é elevada e turmas com altos valores do coeficiente β_{2k} indicam que a leitura tem grande influência no desempenho dos alunos em português.

A etapa seguinte é explicar a variação dos coeficientes de regressão β_{ik} introduzindo a variável explicativa Z associada a turma.

$$\beta_{0k} = \gamma_{00} + \gamma_{01}Z_k + u_{0k}, \quad (1.2)$$

$$\beta_{1k} = \gamma_{10} + \gamma_{11}Z_k + u_{1k}, \quad (1.3)$$

$$\beta_{2k} = \gamma_{20} + \gamma_{21}Z_k + u_{2k}. \quad (1.4)$$

O modelo apresentado na equação 1.2 prediz a média(β_{0k}) do desempenho em português pela experiência do professor(Z). De modo geral, se γ_{01} é positivo então a média do desempenho é maior em turmas com professores mais experientes e se γ_{01} for negativo ocorre o inverso.

A relação entre o desempenho (Y) e o sexo (X_1), representada na equação 1.3, depende da experiência do professor (Z). Se γ_{11} é positivo então o efeito do sexo no desempenho é maior em turmas com professores mais experientes, se γ_{11} é negativo ocorre o oposto.

Analogamente, a relação entre o desempenho (Y) e a Leitura (X_2), representada na equação 1.4, depende da experiência do professor (Z). Se γ_{21} é positivo então o efeito da leitura é maior em turmas com professores mais experientes, se γ_{21} é negativo ocorre

o contrário.

A variável aleatória Z age como uma variável moderadora para a relação entre o desempenho, o sexo e a leitura.

Os resíduos aleatórios no nível superior estão representados por u_{0k}, u_{1k} e u_{2k} , assumem média zero e são independentes dos resíduos e_{ij} associados ao nível inferior. As variâncias são dadas por $\sigma_{u0}^2, \sigma_{u1}^2$ e σ_{u2}^2 e as covariâncias não são nulas.

Substituindo-se as equações 1.2, 1.3 e 1.4 na equação 1.1 pode-se chegar no modelo completo.

$$\begin{aligned} Y_{ik} = & \gamma_{00} + \gamma_{10}X_{1ik} + \gamma_{20}X_{2ik} + \gamma_{01}Z_k + \gamma_{11}X_{1ik}Z_k \\ & + \gamma_{21}X_{2ik}Z_k + u_{1k}X_{1ik} + u_{2k}X_{2ik} + u_{0k} + e_{ik}, \end{aligned} \quad (1.5)$$

No modelo apresentado na equação 1.5 existem duas grandes partes, a primeira com coeficientes fixos (parte determinística) e a segunda com coeficientes aleatórios (parte aleatória).

A parte determinística do modelo definido na equação 1.5 é dada por:

$$\gamma_{00} + \gamma_{10}X_{1ik} + \gamma_{20}X_{2ik} + \gamma_{01}Z_k + \gamma_{11}X_{1ik}Z_k + \gamma_{21}X_{2ik}Z_k, \quad (1.6)$$

A parte aleatória do modelo definido na equação 1.5 é dada por:

$$u_{1k}X_{1ik} + u_{2k}X_{2ik} + u_{0k} + e_{ik}. \quad (1.7)$$

Como consequência da regressão dos coeficientes β_k aparecem termos interativos XZ no modelo. Na parte aleatória aparecem termos uX o que significa que o resíduo vai ser diferente dependendo da variável aleatória X , ou seja, o modelo é heterocedástico, diferentemente do modelo de regressão tradicional no qual se tem a suposição de homocedasticidade.

O modelo de regressão multinível com dois níveis e p variáveis explicativas no nível inferior e q variáveis explicativas no nível superior, conforme Hox (2010) é definido

por:

$$Y_{ij} = \gamma_{00} + \gamma_{p0}X_{pij} + \gamma_{0q}Z_{qj} + \gamma_{pq}Z_{qj}X_{pij} + u_{pj}X_{pij} + u_{0j} + e_{ij}, \quad (1.8)$$

O que claramente possui um número elevado de parâmetros.

1.3 Estimação dos Parâmetros

O método de estimação geralmente utilizado para estimar os parâmetros do modelo é o Método de Máxima Verossimilhança. Ele possui vantagens tais como produzir estimativas que são consistentes e assintoticamente eficientes. Em grandes amostras, estimadores de máxima verossimilhança são geralmente robustos contra violações leves de pressupostos, como ter erros que não são normais.

Duas funções de Máxima Verossimilhança são usadas em modelos de regressão multinível.

Uma é a *Máxima Verossimilhança Completa* (MVC), em que os coeficientes de regressão e os componentes da variância são incluídos na função de verossimilhança. O outro método é *Máxima Verossimilhança Restrita* (MVR), em que apenas os componentes da variância são incluídos na função de máxima verossimilhança, e os coeficientes de regressão são estimados em uma segunda etapa. Os dois métodos produzem estimativas dos parâmetros com erro padrão associado e uma *deviance* do modelo global.

A Máxima Verossimilhança Completa trata os coeficientes de regressão como fixos mas desconhecidos quando os componentes da variância são estimados. A Máxima Verossimilhança Restrita estima os componentes da variância após remover os efeitos fixos do modelo, esses estimadores tem menos viés. Porém, na prática a diferença entre os dois estimadores é pequena.

Outros métodos que tem sido usados na estimação dos parâmetros são Mínimos Quadrados Generalizados, bootstrap e métodos Bayesianos.

1.4 Intervalos de Confiança

Os estimadores dos parâmetros gerados através de estimação por máxima verossimilhança têm seus correspondentes erros padrão associados. Tais valores podem ser utilizados no cálculo da estatística do teste de significância Z, conhecido como Teste de Wald, o qual é dado por:

$$z = \frac{\hat{\beta}_j}{\text{ErroPadrao}(\hat{\beta}_j)}, \quad (1.9)$$

A estatística Z tem distribuição normal padrão ($N(0,1)$) e $J-P-1$ graus de liberdade, com J sendo o número de unidades no segundo nível e P o número de variáveis explicativas. Os erros padrão são assintóticos e portanto são válidos para grandes amostras. Não se sabe precisamente qual tamanho a amostra deverá ter para se obter estimadores precisos. Se o interesse é estimar e interpretar os coeficientes de regressão com P variáveis explicativas é sugerido que se tenha uma amostra de tamanho $104+P$ (regra básica). Se o interesse é interpretar a variância o número sugerido é $50+8P$ observações. Em análise de regressão multinível o tamanho da amostra no nível superior é a quantidade de grupos. (Hox, 2010).

Uma outra alternativa para testar os componentes da variância é proposto por Raudenbush e Bryck (2002):

$$\chi^2 = \sum \frac{(\hat{\beta}_j - \beta)^2}{\hat{V}_j}, \quad (1.10)$$

em que β_j é o estimador de mínimos quadrados ordinários calculados separadamente no grupo j , β é a média dos estimadores e \hat{V}_j é a estimação da variância no grupo j . O número de graus de liberdade é $J-P-1$, sendo J o número de unidades no segundo nível e P o número de variáveis explicativas.

1.5 Comparação de Modelos

A partir da função de máxima verossimilhança é possível calcular a estatística chamada *deviance*, a qual é um indicativo da qualidade do ajuste do modelo. Tal estatística é definida como:

$$deviance = -2 * \ln(Verossimilhanca), \quad (1.11)$$

em que Verossimilhança é a função de verossimilhança. Em geral, modelos com um baixa deviance se ajustam melhor do que modelos com alta deviance.

Em modelos encaixados, ou seja, nos casos em que um modelo é um caso particular de um modelo mais geral, a *deviance* pode ser usada para comparação. A diferença da *deviance* entre os dois modelos segue uma distribuição qui-quadrado com d graus de liberdade, onde d é a diferença entre parâmetros dos dois modelos.

Se os modelos não são encaixados pode-se utilizar o Critério de Akaike (AIC), medida baseada na *deviance* d , acrescida de uma penalidade associada ao número de parâmetros q . Esta penalidade tem a finalidade de corrigir um viés resultante da comparação de modelos com diferentes números de parâmetros.

$$AIC = d + 2q, \quad (1.12)$$

O modelo escolhido deve ser o que apresenta menor AIC. Outro índice similar ao AIC é o Critério de Informação Bayesiano(BIC), é dado por:

$$BIC = d + q\ln(N), \quad (1.13)$$

O modelo que melhor se ajusta é o que apresenta menor BIC.

1.6 Seleção de Variáveis

Com o intuito de selecionar as variáveis presentes no modelo, pode-se proceder de duas maneiras. No primeiro caso o modelo inclui o máximo de efeitos fixos e aleatórios que são considerados no modelo. Normalmente esse procedimento é feito em dois passos. Primeiramente seleciona-se os efeitos fixos e as possíveis interações, em seguida são removidos os efeitos que não são significativos. O segundo passo consiste em construir uma estrutura aleatória rica, em seguida é removido os efeitos que não são significativos.

Tal procedimento possui algumas desvantagens, pois ao se começar com um modelo maior e mais complicado alguns problemas podem ocorrer, como não convergir.

No segundo caso se começa com um modelo mais simples, adiciona-se os parâmetros passo a passo, testando a significância após esses parâmetros terem sido adicionados no modelo. Normalmente a parte fixa é a primeira a ser construída no modelo seguida da parte aleatória. Essa técnica costuma selecionar modelos mais simples. Hox (2010) descreve os passos dessa técnica para a construção do modelo:

Passo 1

O modelo é analisado sem nenhuma variável explicativa. O modelo contém apenas o intercepto:

$$Y_{ij} = \gamma_{00} + u_{0j} + e_{ij}. \quad (1.14)$$

No modelo apresentado na equação 1.14, γ_{00} é o intercepto da regressão, u_{0j} e e_{ij} são os resíduos usuais no nível do grupo e no nível individual. O modelo fornece uma estimativa da correlação intraclasse ρ :

$$\rho = \frac{\sigma_{u0}^2}{\sigma_e^2 + \sigma_{u0}^2}, \quad (1.15)$$

σ_{u0}^2 é a variância dos resíduos u_{0j} e σ_e^2 é a variância dos resíduos e_{ij} . Esse modelo também proporciona uma medida de referência do *deviance*, que é uma medida do grau de desajuste do modelo e pode ser usado para comparar modelos, quanto menor

o valor da *deviance* melhor o ajuste.

Passo 2

O modelo é analisado com todas as variáveis explicativas fixas no nível inferior. O que significa que os componentes da variância correspondentes aos coeficientes são fixados em zero. Esse modelo está descrito na equação 1.16:

$$Y_{ij} = \gamma_{00} + \gamma_{p0}X_{pij} + u_{0j} + e_{ij}, \quad (1.16)$$

X_{pij} são as p variáveis explicativas no nível inferior. Nesse passo é estimado a contribuição e cada variável explicativa desse nível.

Passo 3

As variáveis explicativas no nível superior são acrescentadas nesse passo:

$$Y_{ij} = \gamma_{00} + \gamma_{p0}X_{pij} + \gamma_{0q}Z_{qj} + u_{0j} + e_{ij}, \quad (1.17)$$

Z_{qj} são as q as variáveis explicativas no nível superior.

Os modelos apresentados nas equações 1.16 e 1.17 são chamados de modelos de componentes de variância, pois esses modelos decompõem a variância do intercepto em componentes distintos da variância para cada nível hierárquico. Assume-se, nesses modelos, que o intercepto varia entre o nível superior, porém os coeficientes de regressão são considerados fixos.

Passo 4

No quarto passo é avaliado se algum dos coeficiente das variáveis explicativas tem componente da variância entre os grupos significativo. Esse modelo é chamado de *modelo de coeficientes aleatórios*:

$$Y_{ij} = \gamma_{00} + \gamma_{p0}X_{pij} + \gamma_{0q}Z_{qj} + u_{pj}X_{pij} + u_{0j} + e_{ij}, \quad (1.18)$$

u_{pj} são os resíduos no nível superior das variáveis explicativas X_{pij} no nível inferior.

Passo 5

Nessa etapa são adicionadas as interações entre-níveis entre variáveis explicativas no nível superior e as variáveis explicativas no nível inferior que tiveram variância significativa dos coeficientes no passo anterior.

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + \gamma_{01}Z_j + \gamma_{11}X_{ij}Z_j + u_{1j}X_{1ij} + u_{0j} + e_{ij}. \quad (1.19)$$

1.7 Coeficiente de Correlação Intraclass

O coeficiente de correlação intraclass mede ao mesmo tempo a homogeneidade em um mesmo grupo e a heterogeneidade em grupos distintos. Esse coeficiente varia entre 0 e 1. Quanto maior a correlação maior a evidência de que o modelo utilizado deverá ser o modelo de regressão multinível.

Com o intuito de se estimar o coeficiente de correlação intraclass é utilizado o modelo que contém apenas o intercepto, ou seja, não deverá conter nenhuma variável explicativa.

$$Y_{ij} = \gamma_{00} + u_{0j} + e_{ij}, \quad (1.20)$$

O modelo representado na equação 1.20 decompõe a variância em dois tempos independentes:

- σ_e^2 : Variância dos erros no nível inferior e_{ij}
- σ_{u0}^2 : Variância dos erros no nível superior u_{0j}

Dessa forma podemos obter o coeficiente de correlação intraclass ρ por:

$$\rho = \frac{\sigma_{u0}^2}{\sigma_e^2 + \sigma_{u0}^2}. \quad (1.21)$$

A correlação intraclass ρ representa a proporção da variância explicada pela estrutura hierárquica.

1.8 Coeficiente de Determinação R^2

Em modelos de regressão tradicionais quando deseja-se verificar a proporção da variância modelada pelas variáveis explicativas é utilizado o coeficiente de determinação R^2 .

Em modelos de regressão multinível existem alguns problemas. Primeiro, existe uma grande quantidade de variância não explicada em diversos níveis. Além disso, se existem coeficientes aleatórios o modelo se torna consideravelmente complexo e conceito da variância explicada não é único.

Existem algumas abordagens que indicam a qualidade de ajuste do modelo em regressão multinível. Para se calcular a estatística análoga ao coeficiente de determinação R^2 , no nível inferior pode-se utilizar:

$$R_1^2 = \frac{\sigma_{e|b}^2 - \sigma_{e|m}^2}{\sigma_{e|b}^2}, \quad (1.22)$$

em que $\sigma_{e|b}^2$ é a variância dos resíduos no modelo apenas com o intercepto do nível mais baixo e $\sigma_{e|m}^2$ é a variação do resíduo no nível mais baixo a ser comparado.

A proporção da variância a ser explicada no segundo nível é dada por:

$$R_2^2 = \frac{\sigma_{u_0|b}^2 - \sigma_{u_0|m}^2}{\sigma_{u_0|b}^2}, \quad (1.23)$$

em que $\sigma_{u_0|b}^2$ é a variância dos resíduos no modelo apenas com o intercepto do nível mais alto e $\sigma_{u_0|m}^2$ é a variação do resíduo no nível mais alto a ser comparado.

A interpretação é análoga ao modelo de regressão tradicional.

1.9 Análise do Resíduo

Análise do resíduo é um método utilizado em regressão multinível para verificação das suposições do modelo, tais como normalidade e linearidade. Fazer essa checagem é de extrema importância visto que modelos podem ser complicados. Nesses casos tem-se um resíduo para cada efeito aleatório no modelo.

- Gráficos dos Resíduos

O modelo apresentado na equação 1.24 representa uma versão possível do exemplo dado na seção 1.2. Esse é um modelo multinível sem a interação entre-nível.

$$Y_{ik} = \gamma_{00} + \gamma_{10}X_{1ik} + \gamma_{20}X_{2ik} + \gamma_{01}Z_k + u_{2k}X_{2ik} + u_{0k} + e_{ik} \quad (1.24)$$

Nesse modelo existem três resíduos dos erros, u_{2k} , u_{0k} e e_{ik} . O box-plot é suficiente para identificar a presença de pontos discrepantes. Em análise e regressão multinível é assumido que a variância do resíduo dos erros é a mesma em todos os grupos.

Para testar a suposição de normalidade pode-se plotar o gráfico dos resíduos padronizados versus escores normais. Se os resíduos tiverem distribuição normal então o gráfico deve apresentar uma linha reta na diagonal.

O gráfico de dispersão dos resíduos contra os valores preditos fornece informações sobre possíveis falhas na suposição de normalidade, linearidade e heterocedasticidade. Se as suposições não são violadas, os pontos devem estar dispersos de maneira aleatório em torno do zero, como mostra as Figuras 1.3 e 1.4.

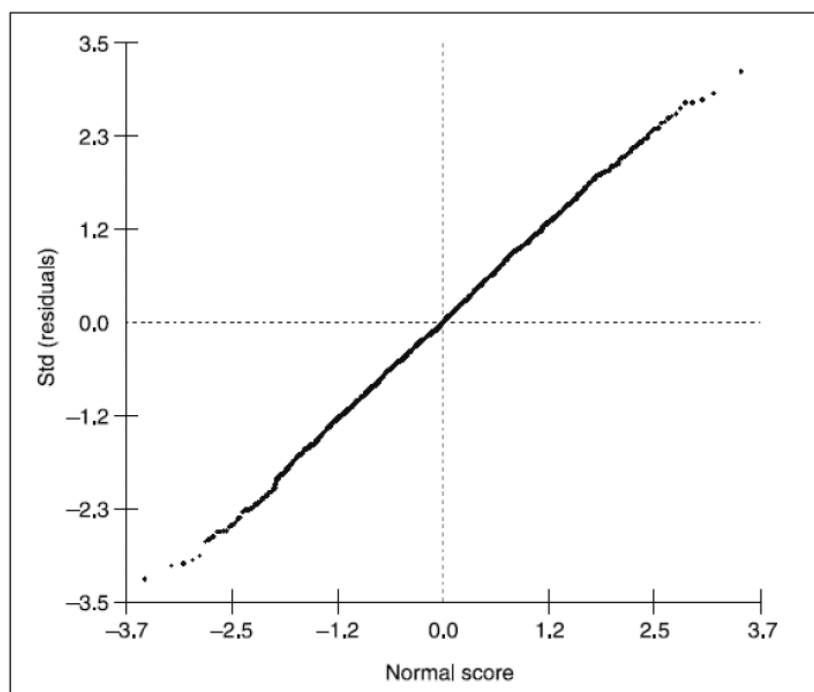


Figura 1.3: Gráfico dos Resíduos Padronizados versus escores normais

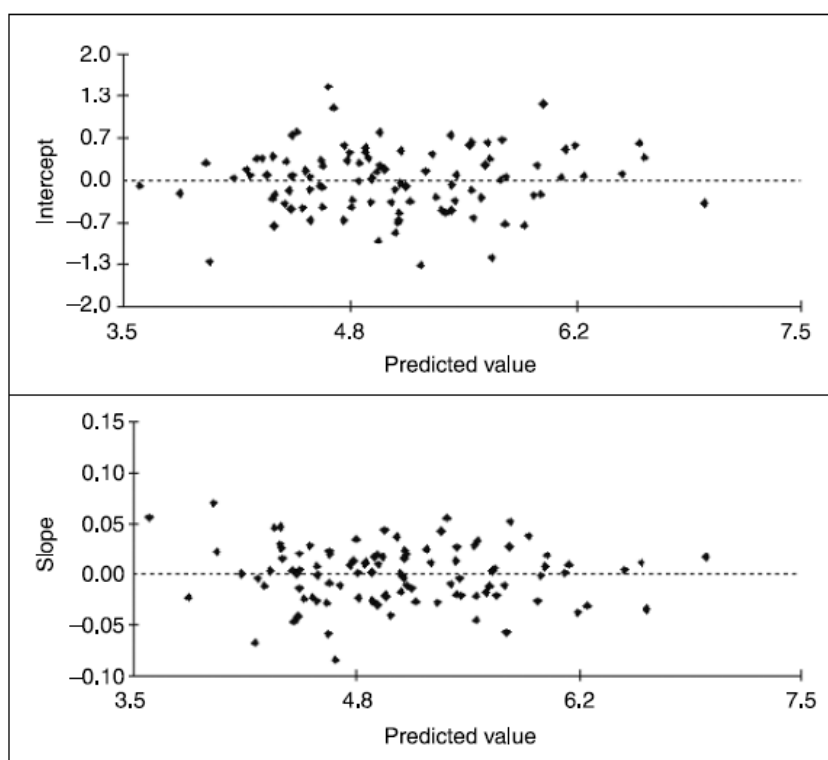


Figura 1.4: Gráfico dos resíduos versus a Variável Preditora

Capítulo 2

Material e Método

2.1 Introdução

Um dos grandes desafios atualmente no Brasil está relacionado com questões referentes a melhoria na qualidade da educação. Dentro do contexto educacional uma das preocupações dos educadores está em detectar quais práticas contribuem para alcançar resultados almejados e quais não. Com o intuito de se aumentar o nível do desempenho escolar dos estudantes é importante conhecer e avaliar a Educação no Brasil.

O Sistema de Avaliação da Educação Básica ,criado em 1990 e coordenado pelo INEP tem como principal objetivo : "avaliar a Educação Básica brasileira e contribuir para a melhoria de sua qualidade e para a universalização do acesso à escola, oferecendo subsídios concretos para a formulação, reformulação e o monitoramento das políticas públicas voltadas para a Educação Básica"(INEP,2015).

O SAEB é aplicado bianualmente para alunos matriculados no 5º e 9º ano do ensino fundamental e no 3º ano do ensino médio em escolas das redes públicas e privadas de todas as Unidades da Federação (Estados e o Distrito Federal).O SAEB é composto por provas e questionários contextuais. As provas avaliam o desempenho escolar dos alunos participantes em Língua Portuguesa e Matemática. Os questionários contextuais são respondidos pelos alunos, diretores, professores e um aplicador responde questões referentes a infraestrutura da escola.

A população de referência do SAEB 2013 são todos os alunos dos 5º e 9º ano do Ensino Fundamental das escolas públicas com no mínimo 20 alunos matriculados, para as escolas com 10 a 19 alunos são coletadas amostras. Também são coletadas amostras

para os alunos dos 5º e 9º do Ensino Fundamental das escolas particulares e alunos do 3º ano do Ensino Médio e do 3º ano do Ensino Médio integrado das escolas públicas e particulares. A população de referência do SAEB 2013 no Brasil é apresentada na Tabela 2.1.

Tabela 2.1: População de referência no Brasil do SAEB 2013

Série	Alunos	Turmas	Escolas
5º Ano do Ensino Fundamental	3.084.936	127.147	65.606
9º Ano do Ensino Fundamental	2.987.772	105.482	49.241
3º Ano do Ensino Médio	2.130.978	68.770	23.914
3º Ano do Ensino Médio Integrado	82.085	3.495	991
Total	8.285.771	304.894	88.720

Fonte: Microdados do SAEB 2013, INEP (2015).

Nesse trabalho o interesse está em verificar o desempenho dos estudantes do Distrito Federal e em particular o rendimento dos alunos do 3º ano do ensino médio. Ao realizar esse estudo observa-se que algumas características estão relacionados aos alunos que por sua vez estão inseridos em escolas que podem desempenhar um papel no rendimento desses estudantes. Como a população estudada está estruturada de forma hierárquica será utilizada modelagem multinível .

Nesse contexto, o interesse do estudo realizado consiste em identificar os fatores associados ao desempenho escolar na prova de Matemática do SAEB 2013 para os estudantes do 3º ano do Ensino Médio do Distrito Federal.

2.2 Material

Os dados utilizados nesse estudo são referentes aos questionários e a prova de matemática das alunos do 3º do Ensino Médio que participaram do SAEB 2013 no Distrito Federal, os quais estão disponibilizados no site do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP).

Foram considerados os alunos do 3º ano do ensino médio regular em área urbana do DF e os dados dos alunos que estavam declarados no Censo Escolar 2012, visto que o Inep apenas considera esses estudantes no cálculo do resultado das escolas. E finalmente, foram desconsiderados os alunos que responderam menos de três itens no teste de proficiência em matemática.

O questionário respondido pelos alunos era composto de 60 questões referentes ao ambiente familiar, hábitos de estudo, motivação, trajetória escolar. O questionário referente aos professores de língua portuguesa e matemática contém 125 questões sobre a formação profissional, nível socioeconômico e cultural, práticas pedagógicas, entre outros. E por fim um aplicador preenche um questionário com 74 questões com informações relacionadas a infraestrutura da escola.

2.3 Método

Os alunos foram separados em classes socioeconômicas de acordo com o Critério de Classificação Econômica Brasil válido a partir de 01/01/2014, criado pela Associação Brasileira de Empresas de Pesquisa. O critério leva em consideração a posse dos seguintes itens: televisão em cores, rádio, banheiro, automóvel, empregada mensalista, máquina de lavar, videocassete e/ou DVD, geladeira, freezer e por fim o grau de instrução do chefe da família. No questionário contextual aplicado aos alunos consta de todas essas informações, apenas não é especificado quem é o chefe da família, porém existe a informação da escolaridade dos pais e se o aluno mora com estes. No cálculo do critério foi considerado a escolaridade do pai ou responsáveis cujo o aluno mora, caso o aluno more com o pai e a mãe foi considerado a maior escolaridade.

O Critério Brasil separa em 8 classes sendo estas respectivamente A1, A2, B1, B2, C1, C2, D e E. Este critério retorna uma estimativa da renda média bruta mensal familiar, algumas classes são agrupadas devido a proximidade dos valores como pode ser observado na Tabela 2.2.

Tabela 2.2: Renda média bruta familiar no mês em Reais por classe

Classes	Renda Média Bruta familiar no mês em Reais
Classe A (A1 e A2)	11.037
Classe B1	6.006
Classe B2	3.118
Classe C1	1.865
Classe C2	1.277
Classe DE	895

Fonte: LSE 2012 Ibope Media

Na primeira etapa foi feita uma análise descritiva com as principais características dos alunos e das escolas. Em seguida foi feita uma análise bivariada, verificando de

maneira exploratória quais variáveis influenciam no desempenho dos alunos na prova de matemática do SAEB. Para as variáveis que possuem dois grupos foi realizado o teste t para amostras independentes. Na comparação de médias múltiplas foi feita uma Análise de Variância (ANOVA). Outra análise bivariada foi feita através de boxplots com o intuito de verificar graficamente os resultados encontrados nos testes de comparação.

Na modelagem multinível foram consideradas as variáveis significativas da etapa anterior. As variáveis referentes aos professores foram agregadas a escola, pois essa foi considerada como segundo nível.

A modelagem multinível foi realizada no *software* estatístico R (R Core Team, 2015) utilizando-se o pacote *lme4*, desenvolvido por Bates et al. (2015). Na construção do modelo foram seguidos os passos propostos por Hox (2010). E por fim, foi realizado a análise do modelo final.

Capítulo 3

Resultados e Discussões

3.1 Perfil dos Alunos

Nesse estudo foi analisado o perfil de 3.041 estudantes do 3º ano do ensino médio que participaram da avaliação do SAEB em 2013 no Distrito Federal.

O questionário respondidos pelos alunos contém 60 perguntas. Os itens constantes no questionário são referentes aos seguintes categorias:

- **Sociodemográficas:** sexo, raça, idade, bens da casa do aluno;
- **Familiares:** se mora com a mãe (mulher responsável) e/ou pai (homem responsável), escolaridade e hábitos de leitura dos pais, incentivo dos pais em relação aos estudos, acompanhamento escolar pelos responsáveis;
- **Cognitivos:** Hábitos de leitura dos alunos, tempo gasto com atividades extraclasse, se o aluno trabalha, se o aluno é repetente, se o aluno já abandonou a escola, se aluno gosta de Língua Portuguesa e/ou Matemática, relação dever de casa/material consultado.

De acordo com a legislação que organiza a oferta de ensino no país (Lei 9.394/1996), o adolescente deve estar matriculado no ensino médio na faixa etária dos 15 aos 17 anos. As provas do SAEB foram aplicadas no mês de novembro e portanto os alunos com até 18 anos estão na idade própria para a série analisada. Pode-se observar na Tabela 3.1 que 82% dos alunos estão na idade adequada.

Como pode ser observado na Tabela 3.2 existe uma diferença na distribuição dos alunos de acordo com a raça/cor entre escolas públicas e particulares. A quantidade

Tabela 3.1: Distribuição de frequência dos alunos do 3º ano do ensino médio do Distrito Federal por idade

Idade	Frequência n = 3.041
Até 16 anos	4%
17 anos	48%
18 anos	30%
19 anos	6%
20 anos	2%
21 anos	11%

Fonte:Microdados do SAEB 2013, INEP 2015.

de estudantes brancos das escolas privadas é praticamente o dobro dos alunos brancos das escolas públicas. Em contrapartida 67,7% dos estudantes das escolas públicas são negros (pardos e pretos), nas escolas privadas este valor representa 42,8%.

Tabela 3.2: Distribuição de frequência do sexo e raça dos alunos do 3º ano do ensino médio do Distrito Federal por dependência administrativa da escola

Características	Pública n = 1.571	Particular n = 1.470	Total n = 3.041
Sexo			
Masculino	42,5%	49,0%	45,7%
Femino	57,5%	51,0%	54,3%
Raça/ Cor			
Branco(a)	24,2%	48,2%	35,8%
Pardo(a)	51,5%	35,7%	43,8%
Preto(a)	16,2%	7,1%	11,8%
Amarelo(a)	4,1%	4,8%	4,5%
Indígena	1,9%	0,4%	1,2%
Não Sabe	2,1%	3,8%	3,0%

Fonte:Microdados do SAEB 2013, INEP 2015.

Conforme as informações apresentadas nas Tabelas 3.3 e 3.4 pode-se perceber que a escolaridade dos pais dos alunos de acordo com a dependência administrativa da escola diferem. A maioria das mães e pais dos alunos das escolas particulares têm ensino superior completo (62,2% das mães e 59,3% dos pais). A quantidade de pais e mães com ensino fundamental (1 ou 2) incompleto se destaca entre os alunos das escolas públicas, 29,6% das mães de alunos estudantes de escolas públicas têm ensino fundamental incompleto contra 2,7% nas escolas particulares.

Tabela 3.3: Distribuição de frequência da escolaridade das mães dos alunos do 3º ano do ensino médio do Distrito Federal por dependência administrativa da escola

Escolaridade da Mãe	Pública n = 1.571	Particular n = 1.470	Total n = 3.041
Ensino Fundamental 1 Incompleto	13,5%	0,8%	7,3%
Ensino Fundamental 2 Incompleto	16,1%	1,9%	9,2%
Ensino Médio Incompleto	14,1%	4,6%	9,5%
Ensino Superior Incompleto	36,9%	28,1%	32,6%
Ensino Superior Completo	14,1%	62,2%	37,5%
Não Souberam Informar	5,3%	2,4%	3,9%

Fonte:Microdados do SAEB 2013, INEP(2015)

Tabela 3.4: Distribuição de frequência da escolaridade dos pais dos alunos do 3º ano do ensino médio do Distrito Federal por dependência administrativa da escola

Escolaridade do Pai	Pública n = 1.571	Particular n = 1.470	Total n = 3.041
Ensino Fundamental 1 Incompleto	16,8%	0,8%	9,1%
Ensino Fundamental 2 Incompleto	12,3%	2,2%	7,4%
Ensino Médio Incompleto	12,9%	3,3%	8,3%
Ensino Superior Incompleto	33,7%	30,2%	32,0%
Ensino Superior Completo	11,5%	59,3%	34,6%
Não Souberam Informar	12,8%	4,2%	8,6%

Fonte:Microdados do SAEB 2013, INEP (2015)

As informações em relação a escolaridade dos pais foram utilizadas no cálculo do Critério Brasil para a classificação dos níveis socioeconômicos dos estudantes. Com base nos dados apresentados na Tabela 3.5 percebe-se 48,2% estudantes das escolas públicas e particulares são das classes B1 ou B2. Dentre os alunos das escolas particulares 34,4% pertencem a classe A.

Com relação a participação dos pais na vida escolar dos alunos não existe discrepância a respeito do incentivo dos pais, porém é interessante notar que 30,2% dos pais que nunca (ou quase nunca) vão a reunião de pais são dos alunos das escolas particulares, como mostrado na Tabela 2.6. Deve-se ressaltar que a maioria dos alunos que já reprovaram, abandonaram ou concluíram o ensino fundamental na educação de jovens e adultos são oriundos da escola pública, pois este fator pode influenciar na desempenho dos alunos nas provas do SAEB.

Tabela 3.5: Distribuição de frequência do nível socioeconômico dos alunos do 3º ano do ensino médio do Distrito Federal por dependência administrativa da escola

Nível Socioeconômico	Pública n = 1.571	Particular n = 1.470	Total n = 3.041
Classe A (A1 e A2)	2,5%	34,4%	17,9%
Classe B1	10,5%	34,2%	22,0%
Classe B2	30,0%	22,2%	26,2%
Classe C1	30,5%	6,9%	19,1%
Classe C2	18,3%	2,0%	10,4%
Classes D e E	8,2%	0,3%	4,4%

Fonte: Microdados do SAEB 2013, INEP (2015)

Tabela 3.6: Distribuição de frequências da participação dos pais na vida escolar dos alunos do 3º ano do ensino médio do Distrito Federal por dependência administrativa da escola

Participação dos pais	Pública n = 1.571	Particular n = 1.470	Total n = 3.041
Frequência com que os pais vão à reunião de pais			
Sempre ou quase sempre	51,7%	40,0%	46%
De vez em quando	32,4%	29,8%	31,1%
Nunca ou quase nunca	16,0%	30,2%	22,9%
Pais ou responsáveis incentivam a estudar			
Sim	98,0%	98,1%	98,1%
Não	2,0%	1,9%	1,9%
Pais ou responsáveis incentivam você a fazer o dever de casa e/ou os trabalhos da escola			
Sim	90,4%	89,4%	89,9%
Não	9,6%	10,6%	10,1%
Pais ou responsáveis incentivam você a ir a escola e/ou não faltar às aulas			
Sim	96,4%	97,4%	87,6%
Não	3,6%	2,6%	12,4%
Pais ou responsáveis conversam com você sobre o que acontece na escola			
Sim	68,8%	77,9%	96,9%
Não	31,2%	22,1%	3,1%

Fonte: Microdados do SAEB 2013, INEP (2015)

Os hábitos de estudo dos alunos das escolas públicas e particulares não destoam, porém a maior parte dos professores que não passam dever de casa são provenientes das escolas públicas.

Tabela 3.7: Distribuição de frequências do histórico escolar dos alunos do 3º ano do ensino médio do Distrito Federal por dependência administrativa da escola

Histórico Escolar	Pública n = 1.571	Particular n = 1.470	Total n = 3.041
Reprovação			
Sim	38,9%	10,9%	25,3%
Não	61,1%	89,1%	74,7%
Abandono			
Sim	95,0%	98,9%	3,1%
Não	5,0%	1,1%	96,9%
Concluiu o Ensino Fundamental na Educação de Jovens e Adultos(EJA)			
Sim	6,7%	2,7%	4,8%
Não	93,3%	97,3%	95,2%

Fonte:Microdados do SAEB 2013, INEP (2015)

Tabela 3.8: Distribuição de frequência para os hábitos de estudo dos alunos do 3º ano do ensino médio do Distrito Federal e dos professores por dependência administrativa da escola

Dever de Casa	Pública n = 1.571	Particular n = 1.470	Total n = 3.041
Faz o dever de casa de Matemática			
Sempre ou quase sempre	41,8%	45,1%	43,4%
De vez em quando	44,5%	41,0%	42,8%
Nunca ou quase nunca	11,9%	13,2%	12,5%
O professor não passa dever de casa	1,8%	0,8%	1,3%
Consulta caderno para fazer o dever de casa			
Sim	89,2%	91,1%	90,2%
Não	8,1%	7,5%	7,8%
O professor não passa dever de casa	2,7%	1,4%	2%
Consulta livros didáticos para fazer o dever de casa			
Sim	77,0%	84,8%	80,9%
Não	19,6%	14,2%	16,9%
O professor não passa dever de casa	3,4%	1,0%	2,2%
Consulta material complementar para fazer o dever de casa			
Sim	57,0%	60,5%	58,7%
Não	38,8%	37,6%	38,2%
O professor não passa dever de casa	4,2%	1,9%	3,0%
Consulta sites da internet para fazer o dever de casa			
Sim	75,1%	64,4%	69,8%
Não	21,2%	33,7%	27,4%
O professor não passa dever de casa	3,7%	1,9%	2,8%
Professor corrige o dever de casa de Matemática			
Sempre ou quase sempre	74,8%	87,6%	81,0%
De vez em quando	16,6%	10,6%	13,7%
Nunca ou quase nunca	6,0%	1,0%	3,6%
O professor não passa dever de casa	2,6%	0,9%	1,8%

Fonte:Microdados do SAEB 2013, INEP (2015)

3.2 Perfil das Escolas

Neste estudo foram analisadas 52 escolas. Em relação a dependência administrativa elas foram classificadas como pública e particular, sendo 20 escolas públicas com 1.571 alunos e 32 escolas particulares com 1.470 alunos. Na Figura 3.1 é apresentada a

distribuição dos alunos por dependência administrativa das escolas.

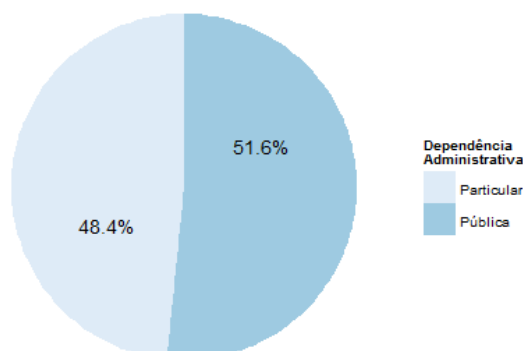


Figura 3.1: Distribuição dos alunos por dependência administrativa das escolas

O perfil dos professores é apresentado na Tabela 3.9. Dentre os professores analisados, 23,2% não possuem licenciatura em matemática. A maior parte dos professores das escolas públicas, 78%, possuem especialização. Apenas 2% dos professores das escolas públicas possuem mestrado, enquanto que no quadro de professores das escolas particulares, 22,2% possuem mestrado. Os professores mais experientes são os das escolas públicas. Somente os professores das escolas particulares trabalham em 3 ou mais escolas e possuem uma carga horária semanal de 39 horas ou menos.

É importante observar que 54,3% dos professores das escolas públicas desenvolveram menos de 80% do conteúdo previsto no ano, em contrapartida, 94,4% dos professores das escolas particulares desenvolveram mais de 80% do conteúdo previsto, o que pode refletir no desempenho dos alunos nas avaliações do SAEB.

Tabela 3.9: Distribuição de frequências do professor por dependência administrativa da escola

Características do Professor	Pública n = 20	Particular n = 32	Total n = 52
Nível de Escolaridade			
Ensino Superior - Licenciatura em Matemática	80,9%	72,0%	76,6%
Ensino Superior - Outra Licenciatura	16,4%	24,4%	20,0%
Ensino Superior - Outras áreas	2,7%	3,6%	3,2%
Curso de Pós-Graduação de Mais Alta Titulação			
Não fez ou não completou	13,8%	35,0%	24,0%
Atualização ou Aperfeiçoamento (mínimo de 180 horas)	6,2%	9,9%	8,0%
Especialização (mínimo de 360 horas)	78,0%	32,9%	56,2%
Mestrado	2,0%	22,2%	11,7%
Área Temática do Curso de Pós-Graduação de Mais Alta Titulação			
Não fez ou não completou	20,1%	33,4%	26,5%
Educação, enfatizando educação matemática	37,1%	27,6%	32,5%
Educação - outras ênfases	29,6%	13,1%	21,6%
Outras áreas que não a Educação	13,2%	25,9%	19,4%
Experiência como Professor			
Menos de 1 ano	4,2%	0,6%	2,5%
1-2 anos	-	5,0%	2,4%
3-5 anos	3,9%	10,1%	6,9%
6-10 anos	16,2%	18,7%	17,4%
11-15 anos	10,8%	26,9%	18,5%
16-20 anos	15,1%	9,3%	12,3%
Mais de 20 anos	49,8%	29,4%	40,0%
Quantidade de Escolas Trabalhadas			
Em 1 escolas	75,1%	49,6%	62,6%
Em 2 escolas	24,9%	31,0%	27,9%
Em 3 escolas	-	15,3%	7,5%
Em 4 ou mais escolas	-	4,1%	2,0%
Carga Horária Semanal			
Mais de 40 horas	45,9%	62,0%	54,1%
40 horas	54,1%	3,9%	28,4%
De 20 a 39 horas	-	32,3%	16,5%
Menos de 20 horas	-	1,8%	0,9%
Horas Semanais Dedicadas às Atividades Extraclasse			
Nenhuma	-	12,9%	6,2%
Menos de 1/3 da carga horária	48,2%	42,0%	45,2%
1/3 da carga horária	33,4%	16,3%	25,1%
Mais de 1/3 da carga horária	18,4%	28,8%	23,4%
Utiliza Programas/Aplicativos Pedagógicos de Computador para Fins Pedagógicos			
Nunca	18,0%	17,3%	17,6%
De vez em quando	60,3%	53,7%	57,2%
Sempre ou quase sempre	21,7%	29,0%	25,2%
Porcentagem do Conteúdo Previsto Desenvolvido			
De 40% a menos de 60%	8,7%	-	4,5%
De 60% a menos de 80%	45,6%	5,6%	26,4%
80% ou mais	45,7%	94,4%	69,1%

Fonte: Microdados do SAEB 2013, INEP (2015)

3.3 Identificação de Possíveis Variáveis Explicativas

Os resultados da análise bivariada para identificar de forma exploratória as variáveis que influenciam no desempenho dos alunos na prova de matemática do SAEB são apresentados nas Tabelas 3.10 e 3.11.

Tabela 3.10: Média do rendimento na prova de matemática do SAEB 2013 para o 3º ano do Ensino Médio no Distrito Federal segundo características dos alunos e Estatísticas dos teste de comparação

Características		Frequência	Média	Estatística	p-valor
Sexo	Masculino	46%	311,54	8,342	<2,2e-16
	Feminino	54%	292,27		
Raça/Cor	Branco(a)	36%	319,74	36,33	<2e-16
	Pardo(a)	44%	291,40		
	Preto(a)	12%	284,16		
	Amarelo(a)	4%	292,01		
	Indígena	1%	257,56		
	Não Sabe	3%	315,99		
Trabalha Fora	Sim	23%	277,08	-12,24	<2,2e-16
	Não	77%	307,92		
Reprovado	Não	75%	314,06	23,25	<2,2e-16
	Sim	25%	261,69		
Abandonou	Não	97%	301,77	4,85	4,701e-06
	Sim	3%	268,76		
Ensino Fundamental EJA	Sim	5%	265,39	-6,93	1,016e-10
	Não	95%	303,01		
Faz o Dever de Casa	Sempre	43%	310,46	18,89	3,92e-12
	Às vezes	43%	295,42		
	Nunca	13%	295,42		
	O Professor Não Passa	1%	272,49		
Nível Socioeconômico	A1	2%	348,67	109,2	<2e-16
	A2	16%	343,20		
	B1	22%	323,89		
	B2	26%	291,72		
	C1	19%	274,27		
	C2	10%	267,83		
	D e E	4%	253,72		
Pais ou responsáveis incentivam a estudar	Sim	98%	301,28	0,72	0,47
	Não	2%	293,97		
Pais ou responsáveis incentivam fazer o dever de casa e/ou os trabalhos da escola	Sim	89,9%	300,35	-0,88	0,38
	Não	10,1%	304,00		
Pais ou responsáveis incentivam ir a escola e/ou não faltar às aulas	Sim	98%	301,62	0,72	0,47
	Não	2%	275,45		

A Tabela 3.10 apresenta os resultados referentes aos alunos. Nessa primeira análise, as únicas variáveis que não foram significativas são as relacionadas ao incentivo dos pais na vida escolar dos alunos. As variáveis que foram significativas nos testes serão consideradas como possíveis variáveis explicativas no modelo multinível, a qual incluirá a variabilidade das escolas e a não independência dos dados.

A Tabela 3.11 indica que há diferença significativa nas variáveis consideradas no nível da escola e portanto serão consideradas como possíveis variáveis explicativas na modelagem multinível.

Tabela 3.11: Características das escolas e Estatísticas dos teste do SAEB 2013 para o 3º ano do Ensino Médio

Características		Frequência	Média	Estatística	p-valor
Dependência Administrativa	Pública	52%	267,39	-35,25	<2,2e-16
	Privada	48%	336,22		
Experiência	Até 2 anos	5%	259,94	48,45	<2e-16
	3 a 10 anos	24%	292,56		
	Mais de 10 anos	71%	305,05		
Quantidade de Escolas	Até 2	91%	297,62	-10,89	<2,2e-16
	3 ou Mais	9%	335,28		
Utiliza Programas/Aplicativos Pedagógicos de Computador	Nunca/ Às Vezes	75%	318,98	-9,79	<2,2e-16
	Sempre	25%	292,80		
Porcentagem do Conteúdo Previsto Desenvolvido no Ano	Menos de 80%	31%	274,33	-16,54	<2,2e-16
	80% ou Mais	69%	312,09		

Os gráficos das variáveis referentes aos alunos são apresentados nas Figuras 3.2 e 3.3. Em geral, os alunos do sexo masculino possuem uma média superior aos do sexo feminino. Os alunos brancos se sobressaem na média obtida no SAEB quando comparados as demais etnias. Quanto maior o nível socioeconômico melhor o desempenho dos estudantes na prova. E, nota-se que os alunos que trabalham fora tem notas menores se comparados aos que não o fazem.

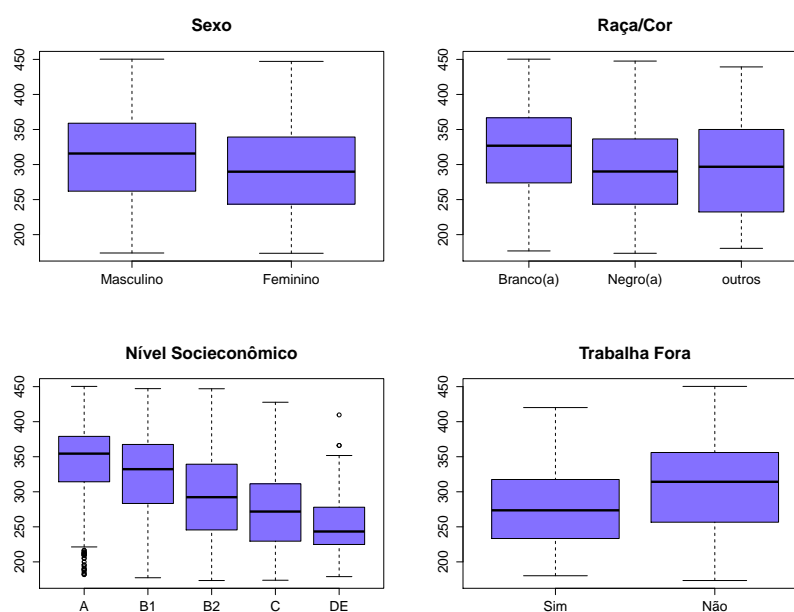


Figura 3.2: Distribuição das notas em Matemática por sexo, raça, nível socioeconômico e se o aluno trabalha

Em relação ao histórico escolar dos alunos percebe-se que os alunos que nunca reprovaram têm média inferior aos demais alunos. Se o aluno já abandonou a escola, ele possui notas menores que os nunca abandonaram. Os alunos que não concluíram o ensino fundamental na educação de jovens e adultos têm melhor desempenho na prova de matemática. E por fim, os alunos que sempre fazem o dever de casa possuem notas melhores que os demais.

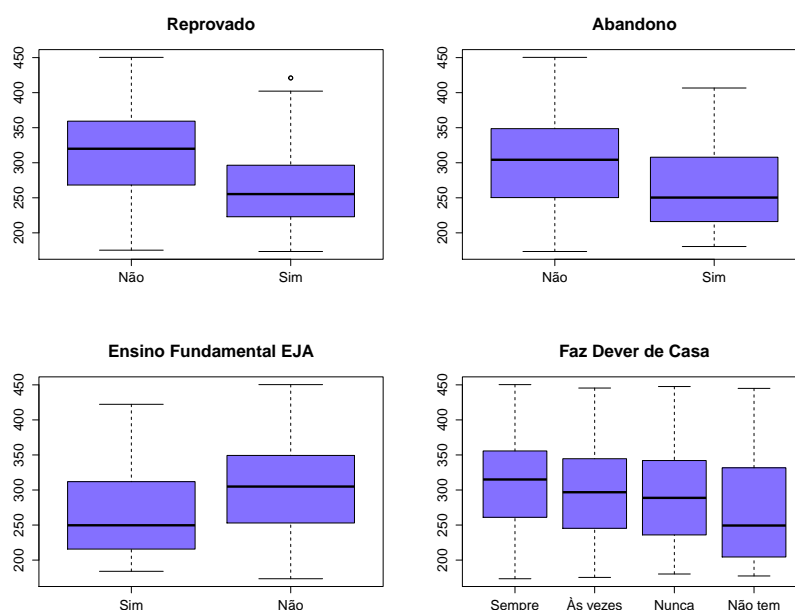


Figura 3.3: Distribuição das notas em Matemática considerando se o aluno já reprovou, abandonou os estudos, concluiu o ensino fundamental no EJA e se faz o dever de casa de matemática

É notável a diferença na distribuição das notas segundo a dependência administrativa da escola. Pode-se também perceber que a nota varia conforme a quantidade de escolas que o professor trabalha, o tempo de experiência do professor, se o professor utiliza programas e aplicativos pedagógicos no computador, assim como varia de acordo com a porcentagem desenvolvida do conteúdo previsto no ano, como mostrado na Figura 3.4.

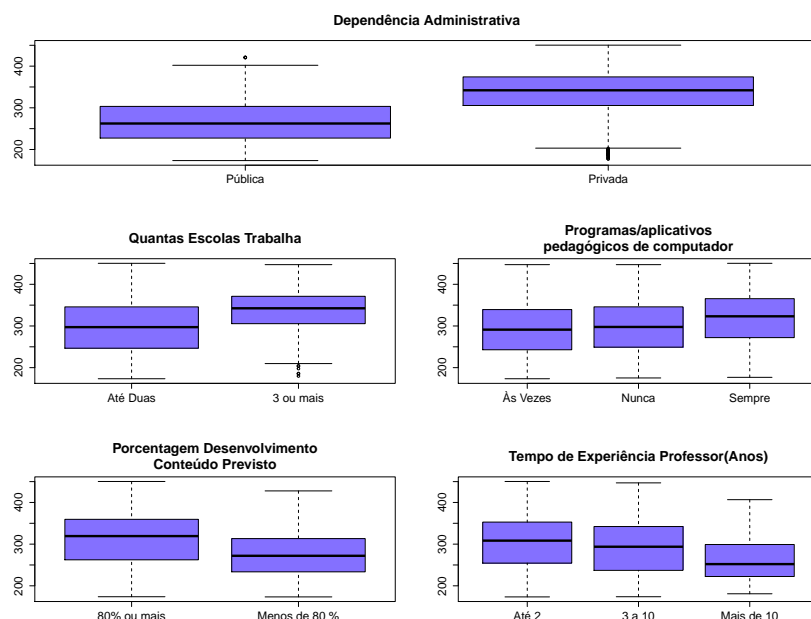


Figura 3.4: Distribuição das notas em Matemática pelas características da escola

3.4 Modelagem

Com base nas possíveis variáveis explicativas definidas na análise bivariada será feita a modelagem, sendo a variável resposta o desempenho dos alunos na prova de matemática do SAEB. A escala escolhida foi a estabelecida pelo SAEB com média 250 e desvio padrão 50, que permite através dos seus valores identificar a capacidade esperada do aluno numa escala de proficiência divulgada pelo SAEB (vide anexo- Escala de Proficiência de Matemática 3ª Série de Ensino Médio). Essa escala assume valores interpretáveis de proficiência entre 225 e 475.

Ajustou-se inicialmente o modelo nulo, que contém apenas o intercepto, e chegou-se ao modelo final seguindo os passos descritos na seção 1.6 do referencial teórico.

Ao ajustar o modelo nulo foi encontrado para o coeficiente de correlação intraclasse o valor de 0,48, o que sugere que 48% da variância do desempenho escolar em Matemática dos alunos estão relacionados à variabilidade entre as escolas, o que torna necessário o uso de modelos de regressão multinível. O valor do intercepto foi de 304,94 que representa o valor médio da proficiência em Matemática dos alunos do 3º ano do Ensino Médio.

Tabela 3.12: Modelo Nulo

Parâmetro fixo	Estimativa	Erro Padrão	Estatística t	p-valor
Intercepto	304,94	6,18	49,35	<2e-16
Parâmetro Aleatório	Variância	Desvio Padrão		
Intercepto	1.921	43,83		
Resíduo	2.061	45,40		
<i>deviance</i>	32.024,01			
Correlação intraclasses	0,48			

No segundo passo as variáveis explicativas fixas no nível dos alunos são adicionadas. Nessa etapa foram incluídas as seguintes variáveis resultantes da análise descritiva:

- **sexo:** Tendo como referência o aluno ser do sexo masculino;
- **raça:** Tendo como referência a cor branca. Na categoria negro foi considerado os alunos pardos e pretos e em outros estão os alunos amarelos, indígenas e os que não sabem como declarar;
- **Trabalha fora:** Se o aluno trabalha fora (recebendo ou não um salário). Tendo como referência se o aluno não trabalha fora;
- **Abandono:** Se o aluno já abandonou a escola durante o período de aulas e ficou fora da escola o resto do ano. Tendo como referência se o aluno não abandonou a escola;
- **Reprovação:** Tendo como referência o aluno que não reprovou;
- **Ensino Fundamental EJA:** Se o aluno não concluiu o Ensino Fundamental na Educação de Jovens e Adultos(EJA);
- **Nível Socioeconômico:** Tendo como referência a classe A (classes A1 e A2);
- **Dever de casa:** Se o aluno nem sempre faz o dever de casa de matemática.

Pode ser observado na Tabela 3.13 que as variáveis que mais influenciam no desempenho do aluno são o sexo e se o aluno possui alguma reprovação. Ao nível de significância de 0,10 as variáveis raça e abandono não foram significativas.

Tabela 3.13: Modelo com todas as variáveis explicativas do primeiro nível

Parâmetro fixo	Estimativa	Erro Padrão	Estatística t	p-valor
Intercepto	314,02	7,32	42,92	<2e-16
Sexo: Feminino	-16,53	1,68	-9,85	<2e-16
Raça: Negro	-2,89	1,86	-1,55	0,12
Raça: Outros	-3,40	3,18	-1,07	0,28
Trabalha Fora: Sim	-4,62	2,17	-2,13	0,03
Reprovação: Sim	-22,01	2,12	-10,38	<2e-16
Abandono: Sim	2,66	5,07	0,52	0,60
Ensino Fundamental EJA: Não	20,39	4,01	5,09	3,81e-07
Classe B1	1,69	2,72	0,62	0,53
Classe B2	-5,96	2,92	-2,04	0,04
Classe C	-7,17	3,18	-2,25	0,02
Classe DE	-12,11	5,03	-2,41	0,02
Dever de Casa Sempre: Não	-9,00	1,71	-5,28	1,40e-07
Parâmetro Aleatório	Variância	Desvio Padrão		
Intercepto	1.416	37,63		
Resíduo	1.873	43,28		
<i>deviance</i>	29.743,6			

No terceiro passo são acrescentadas as variáveis no nível da escola. Nessa etapa foram selecionadas as seguintes variáveis:

- **Dependência Administrativa:** Se a escola é pública ou particular;
- **Média da Escola :** Média dos rendimentos dos alunos da escola;
- **Tempo Experiência Professor:** Proporção de professores por escola com experiência de até dois anos;
- **Quantas Escolas Professor Trabalha:** Proporção de professores por escola que trabalham em até duas escolas;
- **Professor Utiliza Programas:** Proporção de professores por escola que utilizam programas ou aplicativos pedagógicos de computadores;
- **Porcentagem do Conteúdo Previsto Desenvolvido:** Proporção de professores por escola não desenvolveram 80% do conteúdo previsto no ano.

As variáveis relacionadas aos professores foram agregadas à escola visto que 48% das escolas têm dois professores de matemática, assim como a média da escola foi construída a partir das médias dos alunos. Quando considerada a variabilidade na escola

as variáveis raça e abandono continuam não sendo estatisticamente significativas, mantido tudo mais contante. Em relação as escolas, nenhuma variável com informação relacionada ao professor foi estatisticamente significativa, esse fato pode ter ocorrido porque os dados foram agregados, o que pode causar uma perda de informação, sendo assim a análise estatística pode perder poder.

Tabela 3.14: Modelo com acréscimo das variáveis do segundo nível

Parâmetro fixo	Estimativa	Erro Padrão	Estatística t	p-valor
Intercepto	49,84	10,84	4,60	4,52e-06
Sexo: Feminino	-15,90	1,64	-9,68	<2e-16
Raça: Negro	-2,50	1,83	-1,35	0,18
Raça: Outros	-2,77	3,12	-0,89	0,37
Trabalha Fora: Sim	-6,10	2,11	-2,89	0,004
Reprovação: Sim	-19,87	2,07	-9,60	<2e-16
Abandono: Sim	2,94	5,00	0,59	0,56
Ensino Fundamental EJA: Não	18,68	3,92	4,76	2,02e-06
Classe B1	2,03	2,61	0,78	0,44
Classe B2	-4,62	2,79	-1,66	0,098
Classe C	-4,59	3,04	-1,51	0,13
Classe DE	-8,26	4,88	-1,69	0,09
Dever de Casa Sempre: Não	-8,73	1,65	-5,30	1,25e-07
Dependência Administrativa	-3,35	1,57	-2,14	0,03
Média da Escola	0,93	0,03	28,51	<2e-16
Tempo Experiência Professor	-6,50	4,48	-1,45	0,15
Quantas Escolas Professor Trabalha	0,80	3,26	0,25	0,81
Professor Utiliza Programas	-2,28	2,23	-1,02	0,31
Porcentagem do Conteúdo Previsto Desenvolvido	0,77	2,19	0,35	0,73
Parâmetro Aleatório	Variância	Desvio Padrão		
Intercepto	2.736e-12	1.654e-06		
Resíduo	1.859e+03	4.312e+01		
<i>deviance</i>	29.538,4			

Essas variáveis que não foram significativas nessa etapa foram testadas novamente uma a uma sem considerar as outras variáveis não significativas e os resultados foram semelhantes. Dessa forma é apresentada na Tabela 3.15 apenas as variáveis estatisticamente significativas.

Tabela 3.15: Modelo com as variáveis significativas nos dois níveis

Parâmetro fixo	Estimativa	Erro Padrão	Estatística t	p-valor
Intercepto	40,95	9,19	4,46	8,66e-06
Sexo: Feminino	-16,01	1,63	-9,82	<2e-16
Trabalha Fora: Sim	-6,28	2,08	-3,02	0,00257
Reprovação: Sim	-19,28	2,01	-9,60	<2e-16
Ensino Fundamental EJA: Não	17,38	3,83	4,54	5,87e-06
Classe B1	1,56	2,58	0,60	0,55
Classe B2	-4,98	2,75	-1,81	0,069
Classe C	-5,30	2,97	-1,78	0,074
Classe DE	-7,98	4,83	-1,65	0,099
Dever de Casa Sempre: Não	-8,75	1,63	-5,36	9,08e-08
Dependência Administrativa	-2,63	1,31	-2,00	0,045
Média da Escola	0,93	0,03	31,99	<2e-16
Parâmetro Aleatório	Variância	Desvio Padrão		
Intercepto	0	0,00		
Resíduo	1.864	43,17		
<i>deviance</i>	30.026,3			

No quarto passo é verificado se algum dos coeficientes das variáveis explicativas têm variância significativa entre os grupos. Nessa etapa é utilizado o modelo de coeficientes aleatórios. Para esse passo foram utilizadas as variáveis no nível do aluno exceto as classes socioeconômicas.

Tabela 3.16: Modelo com acréscimo dos efeitos aleatórios

Parâmetro fixo	Estimativa	Erro Padrão	Estatística t	p-valor
Intercepto	32,04	10,21	3,13	0,002
Sexo: Feminino	-16,55	2,09	-7,92	2,76e-10
Trabalha Fora: Sim	-5,7	3,26	-1,75	0,088
Reprovação: Sim	-19,65	2,59	-7,57	2,70e-09
Ensino Fundamental EJA: Não	23,09	6,61	3,49	0,001
Classe B1	2,74	2,55	1,08	0,282
Classe B2	-4,37	2,71	-1,61	0,098
Classe C	-3,95	2,93	-1,34	0,079
Classe DE	-8,36	4,79	-1,74	0,081
Dever de Casa Sempre: Não	-9,07	2,12	-4,29	8,65e-05
Dependência Administrativa	-2,87	1,32	-2,17	0,030
Média da Escola	0,94	0,03	31,76	<2e-16
Parâmetro Aleatório	Variância	Desvio Padrão		
Intercepto	531,48	23,05		
Sexo	73,00	8,54		
Trabalha Fora	230,89	15,19		
Reprovação	98,29	9,91		
Ensino Fundamental EJA	916,97	30,28		
Dever de Casa	75,54	8,69		
Resíduo	1741,11	41,73		
<i>deviance</i>	29.940			

O teste do qui-quadrado baseado na *deviance* foi realizado para comparação dos modelos do passo 3 e do passo 4. A hipótese nula é que o modelo apenas com coeficientes fixos é o mais adequado. Como mostrado na Tabela 3.17, existe evidência para rejeitar a hipótese nula e portanto o modelo de coeficientes aleatórios é mais adequado.

Tabela 3.17: Teste χ^2 dos modelos

Modelo	Deviance	χ^2	p-valor
Modelos com apenas coeficientes fixos	30.026	83,01	1,164e-09
Modelo de coeficientes aleatórios	29.940		

No quinto passo foram testadas, uma a uma, todas as possíveis interações entre as variáveis no nível do aluno e no nível da escola e as únicas interações que foram estatisticamente significativas são as seguintes: se o aluno trabalha e a média da escola, se o aluno já reprovou e a média da escola e se o aluno sempre faz o dever de casa e a dependência administrativa da escola. Em seguida foi feita o teste do qui-quadrado baseado na *deviance* comparando o modelo de coeficientes aleatórios e os modelos com as interações que foram significativas, em todos os casos o modelo de coeficientes aleatórios foi o mais adequado.

Com o intuito de se verificar os pressupostos do modelo foi realizada a análise de resíduos para o modelo apresentado na Tabela 3.16. Como pode ser observado nos gráficos 3.5 e 3.6 não há evidências para rejeitar os pressupostos de normalidade, linearidade e heterocedasticidade.

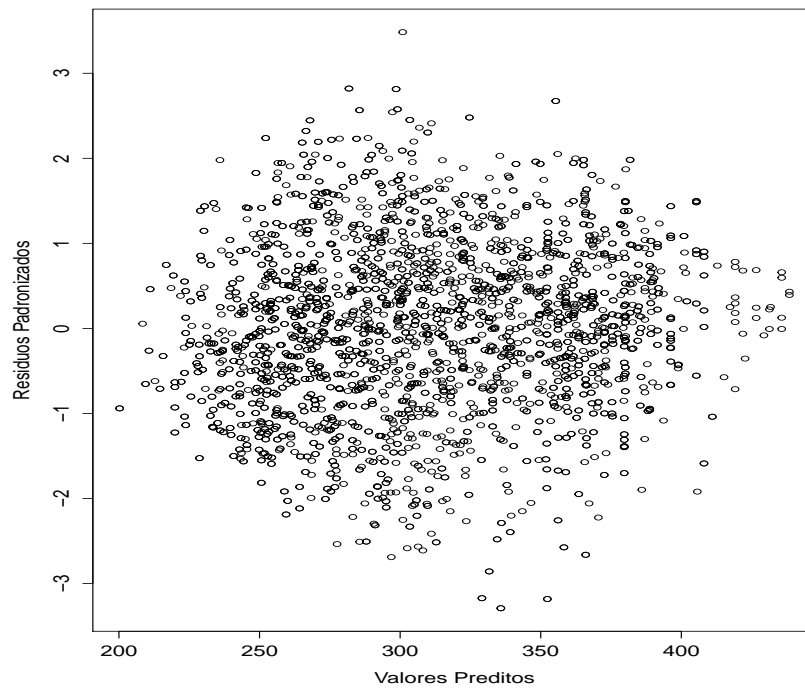


Figura 3.5: Gráfico resíduos padronizados vs. valores preditos para o modelo final

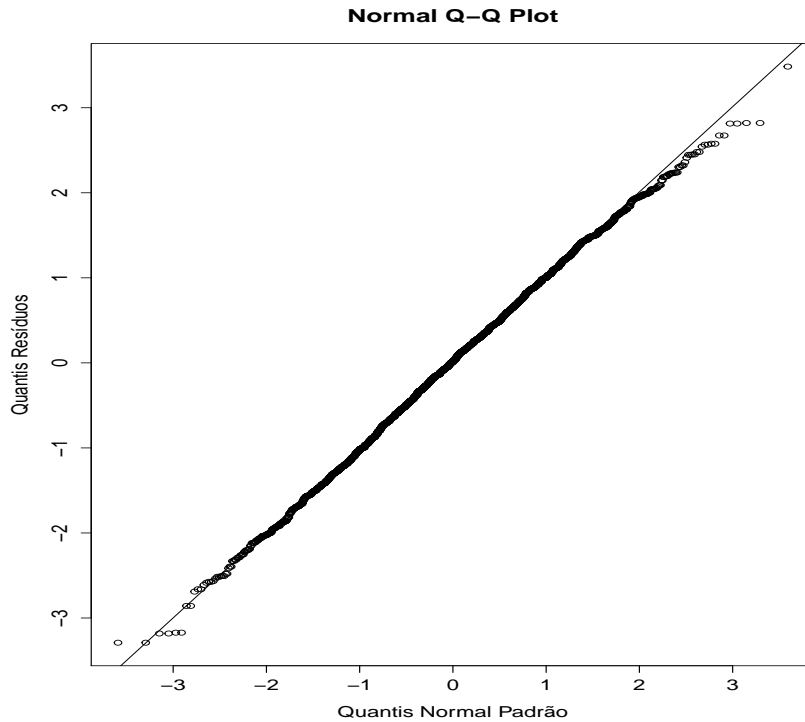


Figura 3.6: Gráfico Quantil-Quantil para o modelo final

Dessa forma, o modelo final é dado da seguinte maneira:

$$\begin{aligned}
 Desempenhoemmatematica_{ij} = & 32,04 - 16,55Sexo_{ij} - 5,7TrabalhaFora_{ij} - 19,65Reprovacao_{ij} \\
 & + 23,09EnsinoFundamentalEJA_{ij} + 2,74ClasseB1_{ij} \\
 & - 4,37ClasseB2_{ij} - 3,95ClasseC_{ij} - 8,36ClasseDE_{ij} \\
 & - 9,07DeverdeCasa_{ij} - 2,87DependnciaAdministrativa_j \\
 & + 0,94MediadaEscola_j
 \end{aligned}$$

Com base nesse modelo, pode-se entender que na prova do SAEB de Matemática para alunos de 3º ano do Ensino Médio no Distrito Federal em 2013:

- Em média, as meninas obtêm 16,55 pontos a menos que os meninos na prova;
- Em média, um aluno que já reprovou alcança 19,65 pontos a menos do que um aluno que nunca reprovou;
- Os alunos que trabalham fora obtêm, em média, 5,7 pontos a menos na prova do que os alunos que não trabalham;

- Os alunos que não concluíram o ensino fundamental na educação de jovens e adultos alcançam, em média, 23,09 pontos a mais do que os que concluíram;
- Os alunos da classe A têm o resultado esperado na prova mais elevado do que as demais classes, exceto se comparado com a classe B1;
- Em média, os alunos que sempre fazem o dever obtêm 9,07 pontos a mais do que os que não o fazem;

Finalmente, é possível estimar a proporção da variância na variável resposta contabilizada em cada nível do modelo. No nível do aluno tem-se que $R_1^2 = 0,1521$ e no nível da escola tem-se $R_2^2 = 0,7236$, ou seja, para este modelo a proporção da variância explicada no nível do aluno é de aproximadamente 15,21% enquanto que no nível da escola é de aproximadamente 72,36%.

Capítulo 4

Conclusão

Com todas as análises feitas, percebe-se que é importante considerar a estrutura hierárquica dos dados analisados nesse trabalho. O coeficiente de correlação intra-classe foi de 0,48, valor elevado, reforçando a importância da modelagem por regressão multinível.

Ao ajustar o modelo descobriu-se que as variáveis associadas ao desempenho escolar em matemática no nível do aluno são o sexo, se o aluno trabalha fora, se houve reprovação, se o aluno conclui o ensino fundamental no EJA, a classe socioeconômica e se o aluno sempre faz o dever de casa. No nível da escola as variáveis dependência administrativa e média da escola são as variáveis que influenciam no desempenho.

Do modelo final pode-se concluir que:

- Os meninos têm notas esperadas mais altas que as meninas;
- Os alunos que não trabalham foram têm notas esperadas mais altas do que os que trabalham;
- Os alunos que já reprovaram têm notas esperadas menor do que os que nunca tiveram reprovação;
- Os alunos da classe A, que possuem uma maior representatividade nas escolas particulares, tendem a apresentar resultados superiores às demais classes, com exceção da classe B1;
- Os alunos que sempre fazem o dever têm notas esperadas maiores do que os que não o fazem;

- A média da escola têm um impacto positivo no desempenho dos alunos.

É interessante observar que 38,9% dos estudantes das escolas públicas possuem alguma reprovação, o que reflete de forma negativa no desempenho dos alunos.

O modelo atendeu as suposições de normalidade e linearidade. A proporção da variância explicada no nível da escola foi de 72,36%, que foi consideravelmente maior do que a explicada no nível do aluno, que foi de 15,21%.

Estudos futuros podem ser feitos observando se as mesmas variáveis estão associadas ao desempenho em português desses alunos. Ou ainda ao desempenho em matemática dos alunos do 5º e 9º ano do Ensino Fundamental.

Anexo

ESCALA DE PROFICIÊNCIA DE MATEMÁTICA 3ª SÉRIE DE ENSINO MÉDIO

Nível*	Descrição do Nível - O estudante provavelmente é capaz de:
Nível 1: 225-250	Tratamento de informações Associar uma tabela de até duas entradas a informações apresentadas textualmente ou em um gráfico de barras ou de linhas.
Nível 2: 250-275	Espaço e forma Reconhecer as coordenadas de pontos representados em um plano cartesiano localizados no primeiro quadrante. Números e operações; álgebra e funções Reconhecer os zeros de uma função dada graficamente. Determinar o valor de uma função afim, dada sua lei de formação. Determinar resultado utilizando o conceito de progressão aritmética. Tratamento de informações Associar um gráfico de setores a dados percentuais apresentados textualmente ou em uma tabela.
Nível 3: 275-300	Números e operações; álgebra e funções Reconhecer o valor máximo de uma função quadrática representada graficamente. Reconhecer, em um gráfico, o intervalo no qual a função assume valor máximo. Determinar, por meio de proporcionalidade, o gráfico de setores que representa uma situação com dados fornecidos textualmente. Determinar o quarto valor em uma relação de proporcionalidade direta a partir de três valores fornecidos em uma situação do cotidiano. Determinar um valor reajustado de uma quantia a partir de seu valor inicial e do percentual de reajuste. Resolver problemas utilizando operações fundamentais com números naturais.
Nível 4: 300-325	Grandezas e medidas Resolver problemas envolvendo área de uma região composta por retângulos a partir de medidas fornecidas em texto e figura. Números e operações; álgebra e funções Reconhecer o gráfico de função a partir de valores fornecidos em um texto. Determinar a lei de formação de uma função linear a partir de dados fornecidos em uma tabela. Determinar a solução de um sistema de duas equações lineares. Determinar um termo de progressão aritmética, dada sua forma geral. Determinar a probabilidade da ocorrência de um evento simples. Resolver problemas utilizando proporcionalidade direta ou inversa, cujos valores devem ser obtidos a partir de operações simples. Resolver problemas de contagem usando princípio multiplicativo.
Nível 5: 325-350	Grandezas e medidas Determinar medidas de segmentos por meio da semelhança entre dois polígonos. Números e operações; álgebra e funções Determinar o valor de variável dependente ou independente de uma função exponencial dada. Determinar o percentual que representa um valor em relação a outro; o valor de uma expressão algébrica. Determinar a solução de um sistema de três equações sendo uma com uma incógnita, outra com duas e a terceira com três incógnitas. Resolver problema envolvendo divisão proporcional do lucro em relação a dois investimentos iniciais diferentes. Resolver operações, além das fundamentais, com números naturais. Resolver a relação linear entre duas variáveis para a determinação de uma delas. Resolver probabilidade de união de eventos. Avaliar o comportamento de uma função representada graficamente, quanto ao seu crescimento.
Nível 6: 350-375	Espaço e forma Reconhecer as coordenadas de pontos representados em um plano cartesiano e localizados em quadrantes diferentes do primeiro. Associar um sólido geométrico simples a uma planificação usual dada. Resolver problemas envolvendo Teorema de Pitágoras, para calcular a medida da hipotenusa de um triângulo pitagórico, a partir de informações apresentadas textualmente e em uma figura. Grandezas e medidas Determinar a razão de semelhança entre as imagens de um mesmo objeto em escalas diferentes. Determinar o volume de um paralelepípedo retângulo, dada sua representação espacial. Números e operações; álgebra e funções Determinar os zeros de uma função quadrática, a partir de sua expressão algébrica. Resolver problemas de porcentagem envolvendo números racionais não inteiros.
Nível 7: 375-400	Espaço e forma Determinar a medida de um dos lados de um triângulo retângulo, por meio de razões trigonométricas, fornecendo ou não as fórmulas. Determinar, com o uso de do teorema de Pitágoras, a medida de um dos catetos de um triângulo retângulo não

	<p>pitagórico.</p> <p>Grandezas e medidas Determinar a área de um polígono não convexo composto por retângulos e triângulos, a partir de informações fornecidas na figura. Resolver problemas por meio de semelhança de triângulos sem apoio de figura. Resolver problemas envolvendo perímetros de triângulos equiláteros que compõem uma figura.</p> <p>Números e operações; álgebra e funções Reconhecer gráfico de função a partir de informações sobre sua variação descritas em um texto. Reconhecer os zeros de uma função quadrática em sua forma fatorada. Reconhecer gráfico de função afim a partir de sua representação algébrica. Reconhecer a equação de uma reta a partir de dois de seus pontos. Reconhecer as raízes de um polinômio apresentado na sua forma fatorada. Determinar os pontos de máximo ou de mínimo a partir do gráfico de uma função. Determinar o valor de uma expressão algébrica envolvendo módulo. Determinar o ponto de interseção de duas retas. Determinar a expressão algébrica que relaciona duas variáveis com valores dados em tabela ou gráfico. Determinar a maior raiz de um polinômio de 2º grau. Resolver problemas para obter valor de variável dependente ou independente de uma função exponencial dada. Resolver problemas que envolvam uma equação de 1º grau que requeira manipulação algébrica. Resolver problemas envolvendo um sistema linear, dadas duas equações a duas incógnitas; usando permutação. Resolver problemas utilizando probabilidade, envolvendo eventos independentes.</p>
Nível 8: 400-425	<p>Espaço e forma Reconhecer a proporcionalidade dos elementos lineares de figuras semelhantes. Determinar uma das medidas de uma figura tridimensional, utilizando o Teorema de Pitágoras. Determinar a equação de uma circunferência, dados o centro e o raio. Determinar a quantidade de faces, vértices e arestas de um poliedro por meio da relação de Euler. Resolver problema envolvendo razões trigonométricas no triângulo retângulo, com apoio de figura. Associar um prisma a uma planificação usual dada.</p> <p>Grandezas e medidas Determinar a área da superfície de uma pirâmide regular. Determinar o volume de um paralelepípedo, dadas suas dimensões em unidades diferentes. Determinar o volume de cilindros.</p> <p>Números e operações; álgebra e funções Reconhecer o gráfico de uma função trigonométrica da forma $y=\sin(x)$. Reconhecer um sistema de equações associado a uma matriz. Determinar a expressão algébrica associada a um dos trechos do gráfico de uma função definida por partes. Determinar o valor máximo de uma função quadrática a partir de sua expressão algébrica e das expressões que determinam as coordenadas do vértice. Determinar a distância entre dois pontos no plano cartesiano. Resolver problema usando arranjo. Resolver problema envolvendo a resolução de uma equação do 2º grau sendo dados seus coeficientes. Interpretar o significado dos coeficientes da equação de uma reta, a partir de sua forma reduzida.</p>
Nível 9: 425-450	<p>Espaço e forma Reconhecer a equação que representa uma circunferência, dentre diversas equações dadas. Determinar o centro e o raio de uma circunferência a partir de sua equação geral. Resolver problemas envolvendo relações métricas em um triângulo retângulo que é parte de uma figura plana dada.</p> <p>Grandezas e medidas Determinar o volume de pirâmides regulares. Resolver problema envolvendo áreas de círculos e polígonos. Resolver problema envolvendo semelhança de triângulos com apoio de figura na qual os dois triângulos apresentam ângulos opostos pelos vértices. Resolver problema envolvendo cálculo de volume de cilindro.</p> <p>Números e operações; álgebra e funções Reconhecer o gráfico de uma função exponencial do tipo $f(x)=10^{x+1}$. Reconhecer o gráfico de uma função logarítmica dada a expressão algébrica da sua função inversa e seu gráfico. Determinar a expressão algébrica correspondente a uma função exponencial, a partir de dados fornecidos em texto ou gráfico. Determinar a inversa de uma função exponencial dada, representativa de uma situação do cotidiano. Determinar inclinação ou coeficiente angular de retas a partir de suas equações. Determinar um polinômio na forma fatorada, dadas as suas raízes.</p>
Nível 10: 450-475	<p>Números e operações; álgebra e funções Determinar a solução de um sistema de três equações lineares, a três incógnitas, apresentado na forma matricial escalonada.</p>

* O intervalo do nível inclui o primeiro ponto e exclui o último ponto.

Referências Bibliográficas

- [1] ASSOCIAÇÃO BRASILEIRA DE EMPRESAS DE PESQUISA. **Critério de classificação econômica Brasil 2014**. São Paulo, Brasil. 2013.
- [2] BATES, D.; MAECHLER, M.; BOLKER, B.; Walker, S. Fitting Linear Mixed-Effects Models Using lme4. **Journal of Statistical Software**, v.67, 2015.
- [3] BATES, D.; MAECHLER, M.; BOLKER, B.; Walker, S. Lme4: Linear Mixed-Effects Models using 'Eigen' and S4. R package version 1.1-10, 2015.
- [4] FERRÃO, M.E; BELTRÃO, K.I.; SANTOS, D.P. O impacto de políticas de não-repetência sobre o aprendizado dos alunos da 4ª série **Pesquisa e Planejamento Econômico**. 2002 Dezembro; 32 (3): 495-514.
- [5] FERRÃO, M.E; BELTRÃO et al. O SAEB ? Sistema Nacional de Avaliação da Educação Básica: objetivos, características e contribuições na investigação da escola eficaz. **Revista Brasileira de Estudos de População**. v.18, p.111-130, 2001.
- [6] GELMAN, A; HILL, J. **Data Analysis Using Regression and Multilevel/Hierarchical Models**. Cambridge University Press, 2007.
- [7] Finch W.H; Bolin, J.E; Kelley, K. **Multilevel Modeling Using R**. CRC Press, 2014.
- [8] Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Microdados da Aneb e da Anresc 2013. [online]. Brasília: Inep, 2015. [citado 2015-05-30]. Disponível em: <<http://portal.inep.gov.br/basicalevantamentos-acessar>>.
- [9] HOX, J. J. **Multilevel analysis: techniques and applications**. Second Edition. Great Britain: Routledge, 2010.

- [10] LAROS, J.A. ; MARCIANO, J.L.M.. Análise multinível aplicada aos dados do nels:88 **Estudos em avaliação educacional** 2008 Maio; 19 (40):263-278.
- [11] LAROS, J.A. ; MARCIANO, J.L.M.; ANDRADE, J.M. Fatores associados ao desempenho escolar em português:um estudo multinível por regiões. **Avaliação e Políticas Públicas em Educação**, v.20, p.623-646, 2012.
- [12] LAROS, J.A. ; MARCIANO, J.L.M.; ANDRADE, J.M.. Fatores que afetam o desempenho na prova de Matemática do SAEB: Um estudo multinível. **Avaliação Psicológica**, v.9 p.173-186, 2010.
- [13] LAROS, J.A. ; MARCIANO, J.L.M.; ANDRADE, J.M. Fatores Associados ao Desempenho Escolar: Estudo Multinível com Dados do SAEB/2001. **Psicologia: Teoria e Pesquisa**, v.33 p.033-042, 2007.
- [14] Monteiro, M.P.S.A. **Modelos de regressão multinível : uma aplicação na educação** 2013. 49p. Dissertação (Mestre em Estatística), Universidade do Minho, 2013.
- [15] Ker,H. W. Application of Hierarchical Linear Models/Linear Mixed-effects Models in School Effectiveness Research. **Universal Journal of Educational Research**. v.2, p. 173-180, 2014.
- [16] R Core Team. **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing, Vienna, Austria, 2015.
- [17] Raudenbush, S. and Bryk, A. . **Hierarchical linear models**. 2nd edition. Sage Publications, 2002.